Supplemental Material 1. Determination of clone library size for capturing breadth of acidobacterial diversity at subdivision level.

Average sequence similarity of subdivisions. It was necessary to calibrate the ability of our small clone libraries to reliably describe our acidobacterial community at the subdivision level. To do this, ca. 664, high-quality nearly full length sequences of subdivision 1, 3, 4, 5, 6, & 8 were downloaded from the ARB-SILVA website (3). These subdivisions were targeted because they are the most common subdivisions captured using the acidobacterial 31F primer (1). Subdivision 8 is not captured with this primer, but sequences were downloaded to use as an outgroup. A neighbor-joining phylogenetic tree was generated in ARB (2) using these sequences and the average sequence similarity of sequences in the respective subdivisions was estimated: subdivision 1 (ca. 91%), 3 (ca. 90%), 4, (ca. 86%), 5 (ca. 91%), and 6 (ca. 91%). Given these average similarities, it appears that the subdivisions level cutoff is ca. 90%. This O.T.U. cutoff was used to test if the sequences in the original clone libraries grouped into their respective subdivision or group using MOTHUR (4).

Assessing Congruency of MOTHUR Groups with the ARB Phylogeny. Preliminary clone libraries from 2 biological replicates (n=50 clones per library, partial sequence ca. 500 bp (31F & 531R)) were used for this analysis. Upon alignment of sequences with SILVA, distance matrices were generated and clustered at 0.15 and above in MOTHUR. In order to assess the congruency of the O.T.U. defined clusters with the ARB phylogeny groupings, we assessed the sequence listed for the respective group (OTU file). We also generated rarefaction curves to determine the number of clone needed to make conclusions at the subdivision level.

These small clone libraries were parsed into their respective subdivisions using ARB phylogeny and reference sequences in the literature (Table 1.1). This clustering was used as the "ground-truth" for the analysis. The 0.15 cut-off was the most accurate cut-off for clustering the subdivisions. The 0.15 cut-off, MOTHUR was ca. 70% accurate. There were minor discrepancies in the clustering, which is most likely due to sequence quality (Table 1.2). Based on these data, we determined that the 0.15 cut-off is at the acidobacterial subdivision level.

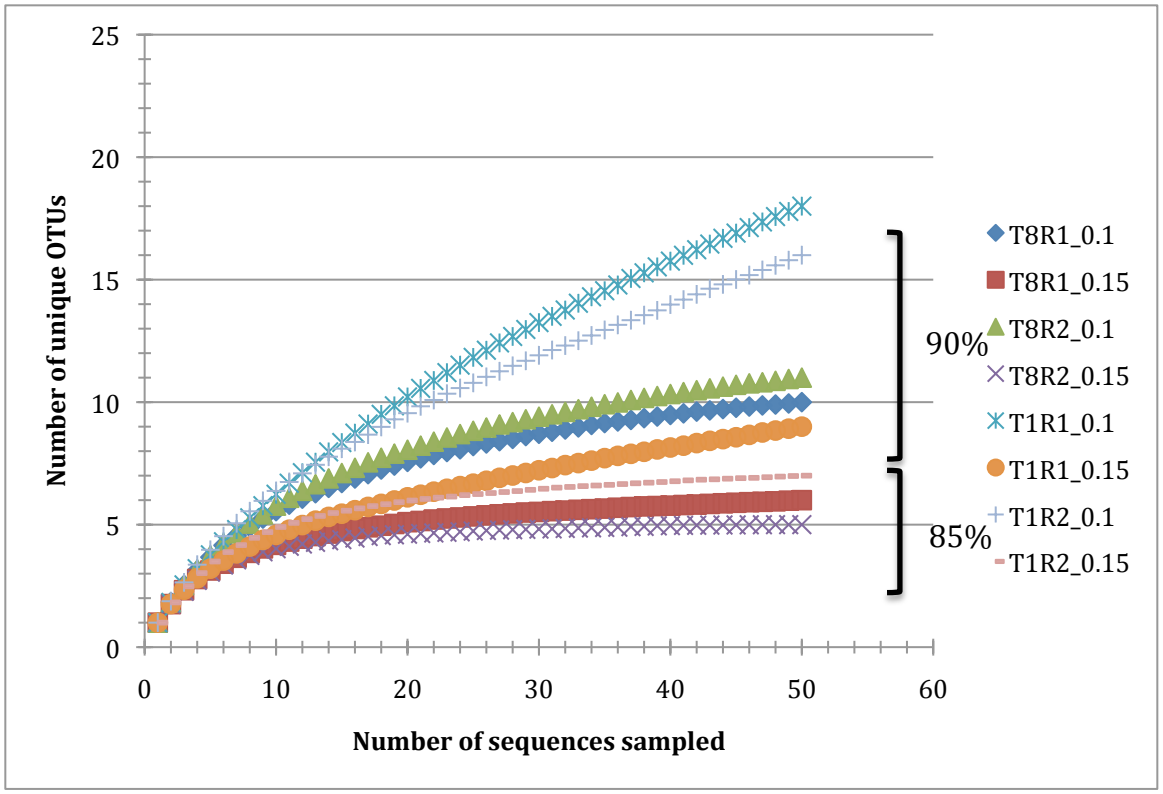
Using these data, we determined the number of clones needed at this cut-off to make conclusions at the subdivision level. At the 85% O.T.U. cut-off, the rarefaction curves level off early in the number of sequences samples – typically ca. 20-30 sequences. One library (orange line) is slightly higher, which is due to the quality of a few sequence reads. Most of the sequences in question between the two methods have long branch lengths in the tree, mostly likely due to poor quality sequencing. Based on rarefaction analysis, one only need to sample ca. 20-30 sequences in order to attain enough coverage to make subdivision level conclusions.

	SG1	**SG3**	**SG4**	**SG5**	**SG6**
T8R1	18	8	3	5	16
T8R2	9	8	9	2	22
T1R1	9	6	26	0	9
T1R2	7	10	19	0	14

Supp. Table 1.1. Subdivision level distribution of the soil samples based on Arb phylogeny using reference sequences, ground truth distribution of small clone libraries. Number of clones per subdivision is depicted in table.

	SG1	SG3	SG4	SG5	SG6
T8R1, 0.15	18	7(8)	3	5	16
T8R2, 0.15	9	8	9	2	22
T1R1, 0.15	8(9)	5(6)	22(26)	0	9
T1R2, 0.15	7	10	16-g3(19) 2-g6	0	8-g1(14) 6-g4

Supp. Table 1.2. MOTHUR clustering of subdivisions for respective libraries at the 0.15 O.T.U. cut-off. Shaded regions depict slight discrepancies between the methods, most likely due to sequence quality. Values listed parenthetically are actual numbers based on Arb phylogeny.



Suppl. Figure 1. Rarefaction analysis of clone libraries at the 0.10 & 0.15 O.T.U. cut-off.

References:

1. **Barns, S. M., S. L. Takala, and C. R. Kuske.** 1999. Wide distribution and diversity of members of the bacterial kingdom *Acidobacterium* in the environment. Appl. Environ. Microbiol. **65:**1731-1737.
2. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. Konig, T. Liss, R. Lussmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K.-H. Schleifer.** 2004. ARB: a software environment for sequence data. Nucl. Acids Res. **32:**1363-1371.
3. **Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glockner.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucl. Acids Res. **35:**7188-7196.
4. **Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber.** 2009. Introducing mothur: open-Source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. **75:**7537-7541.